STIC-Biotech/ChemLib

145042

From:	Chan, Christina
Sent:	Monday, February 14, 2005 9:33 AM
To:	Sullivan, Daniel; STIC-Biotech/ChemLib
Subject:	RE: Rush sequence search 08/876132

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----
From:	Sullivan, Daniel
Sent:	Sunday, February 13, 2005 9:56 AM
To:	Chan, Christina
Subject:	Rush sequence search 08/876132

Hi Chris, Please approve this search as a rush for an after final amended case. Thanks.

Please search for the following in the pending and issued patent databases:

A nucleic acid comprising SEQ ID NO: 1;
A nucleic acid comprising SEQ ID NO: 2;
A nucleic acid encoding SEQ ID NO: 3.

Thank you.

Daniel M. Sullivan
Examiner AU 1636
Remsen Bldg.
Room 2A74

Tel: (571) 272-0779

Mailbox: 2C70

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

On protein - nucleic search, using frame-plus p2n model

Run on: February 15, 2005, 18:22:09; search time 714 seconds
(without alignments)
3065.925 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 19 NPIHHFKPKINQKKLRNKKY .... SQELRLCSFYXRWKMCMC 371

Scoring table: BLOSUM62
Xgapmo 10.0   Xgapext 0.5
Ygapmo 10.0   Ygapext 0.5
Fgapop 6.0   Fgapext 7.0
DelOp   6.0   Delext   7.0

Searched: 5378673 seqs, 2950299984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model -DBV=xln
-OVTOL=1000 -pool_1/USDB07/n132/runat_14022005_12591_2894/app_query-fasta_1.519
-ENDC=0 -DOCKI=0 -UNITS=bit -START=1 -END=-1 -MATRIX=blosum62
-DEL=DOCLO=200 -THR_SCORE=pct -THR_MAX=100
-HW=9 DOCALIGN=200 -QCB=1.0 -XGXN=1.0 -MINLEN=0
-MAXLEN=2000000000 -USERN=USDBR97N132_WCDB_1.0 -ALIGN=1.5
-OKGLOQ -PMAP -LANGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LOKGLOQ -DBV_THR=0 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
-FGAPOP=6 -FGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5

Database:
Published Applications NA.*
1: /cgn2_6/pcdata/i/pubpna/US07_PUBCOMB.seq;*
2: /cgn2_6/pcdata/i/pubpna/USDSU_PUBCOMB.seq;*
3: /cgn2_6/pcdata/i/pubpna/PCT_NEW_PUB.seq;*
4: /cgn2_6/pcdata/i/pubpna/USDSU_PUBCOMB.seq;*
5: /cgn2_6/pcdata/i/pubpna/US06_PUBCOMB.seq;*
6: /cgn2_6/pcdata/i/pubpna/US05_PUBCOMB.seq;*
7: /cgn2_6/pcdata/i/pubpna/US07_NEW_PUB.seq;*
8: /cgn2_6/pcdata/i/pubpna/USTSL_PUBCOMB.seq;*
9: /cgn2_6/pcdata/i/pubpna/US04_PUBCOMB.seq;*
10: /cgn2_6/pcdata/i/pubpna/US03_PUBCOMB.seq;*
11: /cgn2_6/pcdata/i/pubpna/US05_NEW_PUB.seq;*
12: /cgn2_6/pcdata/i/pubpna/USS8_PUBCOMB.seq;*
13: /cgn2_6/pcdata/i/pubpna/US00_PUBCOMB.seq;*
14: /cgn2_6/pcdata/i/pubpna/US99_PUBCOMB.seq;*
15: /cgn2_6/pcdata/i/pubpna/US01_PUBCOMB.seq;*
16: /cgn2_6/pcdata/i/pubpna/US02_PUBCOMB.seq;*
17: /cgn2_6/pcdata/i/pubpna/US10_PUBCOMB.seq;*
18: /cgn2_6/pcdata/i/pubpna/USS0_PUBCOMB.seq;*
19: /cgn2_6/pcdata/i/pubpna/US09_PUBCOMB.seq;*
20: /cgn2_6/pcdata/i/pubpna/USS1_NEW_PUB.seq;*
21: /cgn2_6/pcdata/i/pubpna/US10_NEW_PUB.seq;*
22: /cgn2_6/pcdata/i/pubpna/US09_NEW_PUB.seq;*

SUMMARIES

Pred. No is the number of results predicted by chance to have a
score at least as good as the result being printed,
and is derived by analysis of the total score score distribution.

Result	Score	Query Match	Length	DB ID	Description
1	139.5	7.2	6497	18	US-10-719-993-248	Sequence 248, Appl
2	117.5		2726	19	US-10-383-123-14847	Sequence 20495, A
3	117.5		1762	18	US-10-345-345A-7529	Sequence 7529, Ap
4	117.5		3471	19	US-10-383-123-7330	Sequence 73330, Ap
5	117.5		2873	18	US-10-282-122A-27329	Sequence 27329, A
6	6.0		3071	19	US-10-793-609-602	Sequence 602, Appl
7	6.0		4301	9	US-10-917-800A-1666	Sequence 1666, App
8	6.0		3027	19	US-10-793-609-126	Sequence 126, Appl
9	6.0		6334	19	US-10-283-122A-19657	Sequence 19657, A
10	5.9		7791	18	US-10-311-555-1186	Sequence 1186, App
11	5.9		4840	18	US-10-432-585-6600	Sequence 6600, App
12	5.9		3534	17	US-10-282-122A-16250	Sequence 16250, A
13	5.9		6978	18	US-10-577-134-62	Sequence 62, Appl
14	5.9		7433	18	US-10-723-860-1449	Sequence 1449, Ap
15	5.8		8443	19	US-09-944-865-190	Sequence 190, App
16	5.8		3312	8	US-09-797-385-3	Sequence 3, Appli
17	5.8		3364	9	US-09-797-385-5	Sequence 5, Appli
18	5.8		6272	17	US-10-052-648A-19	Sequence 19, Appl
19	5.8		2256641	18	US-10-470-565-131	Sequence 131, App
20	5.8		6733	17	US-10-222-714-1930	Sequence 1930, Ap
21	5.8		8999	17	US-09-944-865-190	Sequence 190, App
22	5.8		3312	9	US-09-797-385-3	Sequence 7, Appli
23	5.8		3364	9	US-09-797-385-5	Sequence 5, Appli
24	5.8		3011208	17	US-10-398-221-2058	Sequence 70, Appl
25	5.8		2405	9	US-09-748-817-897	Sequence 897, App
26	5.8		1353	18	US-09-851-121-1841	Sequence 2618, Ap
27	5.8		3018	17	US-10-369-493-29748	Sequence 29748, A
28	5.8		2390	17	US-09-706-972	Sequence 634, App
29	5.8		3025	18	US-10-369-493-1841	Sequence 1841, Ap
30	5.8		3025	18	US-10-171-311-471	Sequence 471, App
31	5.7		6031	19	US-10-788-792-39	Sequence 39, Appl
32	5.7		3394	17	US-09-169-487-190	Sequence 7, Appl
33	5.7		5601	17	US-10-424-359-113525	Sequence 113525,
34	5.7		895	17	US-09-693-171144	Sequence 371144, A
35	5.7		805	17	US-10-437-593-10323	Sequence 10323, A
36	5.7		3293	8	US-10-437-593-10323	Sequence 10323, A
37	5.7		2278	18	US-10-543-341-799033	Sequence 799033,
38	5.7		3423	17	US-10-169-493-14081	Sequence 14081, A
39	5.7		3025	18	US-10-169-493-14081	Sequence 14081, A
40	5.7					
41	5.7					
42	5.7					
43	5.7					
44	5.7					
45	5.7					

ALIGNMENTS

RESULT 1
US-10-719-993-248
Patent Application US-10719993
Publication No US2004024584A1
GENERAL INFORMATION:
  TITLE OF INVENTION: Michele et al.
  TITLE OF INVENTION: GENETIC POLYMORPHISM ASSOCIATED WITH
  TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
  CURRENT APPLICATION NUMBER: US10/719,993
  CURRENT FILING DATE: 2003-11-24
  NUMBER OF SEQ ID NOS: 5342
  SOFTWARE (VF SEQ ID NOS:
  SEQ ID NO 248
  LENGTH: 6497
  TYPE: DNA
  ORGANISM: Homo sapiens
US-10-719-993-248

Alignment Scores:

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 18:08:35 ; Search time 229 seconds
(without alignments)
2650.911 Million cell updates/sec

Title: US-08-876-132-3
Perfect Score: 1947
Sequence: 1 NFLHNKRKTKGKKLAKKKY...........SQELLRLCSFVKWVRCNKC 371

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapex 0.5
F9gapop 6.0, F9gapex 0.5
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818133359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Command line parameters:
-MODEL=frame_p2n.model
-Q=/cgn2../USPTO_pgool_1/1/in/9A/app_query.fasta1.519
-DB=/cgn2../in/9A/in.fasta.seq.*
-LOCOPEX=0 -UNITS=bits -MATRIX=blosum62.comp.cel
-LIST=45 -LOCALICE=200 -THR_SCORE=pct -MIN=0 -ALIGN=15
-MODEL=LOCAL.OUTFMT=0 -SCORE=40.0 -STRAND=1
-USER=USER0087K112 WCON 1.1 69 serum 140220065 125915 22812
-HO_MMAP=LACEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-QPAP+EXT=7 -TGHOPT=1 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -DELOP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Database: DDV=xlh

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
c 1	130.5	6.7 1664976 4	US-08-916-421B-1	Sequence 1, Appl1
c 2	130.5	6.5 1664976 4	US-09-692-250-1	Sequence 1, Appl1
c 3	127.5	6.5 43576 4	US-09-676-519-19	Sequence 19, Appl1
c 4	121.5	6.5 43576 4	US-08-545-528D-1	Sequence 1, Appl1
c 5	117	6.5 580073 4	US-08-545-528D-1	Sequence 1, Appl1
c 6	116	6.0 3003 4	US-09-134-001C-2341	Sequence 2341, Appl1
c 7	116	6.0 3881 4	US-09-016-3905	Sequence 3905, Appl1
c 8	114	6.0 3905 4	US-09-134-001C-3905	Sequence 3905, Appl1
9	113.5	5.9 3881 4	US-08-923-992A-5	Sequence 308, Appl1
10	113.5	5.9 3312 4	US-08-923-992A-3	Sequence 308, Appl1
11	113.5	5.8 3384 4	US-09-166-350-27	Sequence 27, Appl1
12	113	5.8 6773 4	US-09-166-350-27	Sequence 27, Appl1
13	112.5	5.8 3025 4	US-08-917-254-19	Sequence 19, Appl1
15	112.5	5.8 3763 4	US-08-916-421B-1	Sequence 1, Appl1
15	112.5	5.7 1664976 4	US-07-792-865D-1	Sequence 7, Appl1
18	111.5	5.7 2617 4	US-08-792-009-1	Sequence 1, Appl1
19	111.5	5.7 2617 4	US-08-714-481-8	Sequence 8, Appl1
20	111.5	5.7 3730 4	US-08-242-932-8	Sequence 8, Appl1
20	111.5	5.7 3730 4	US-08-440-024-1	Sequence 1, Appl1
23	111.5	5.7 3730 4	DTC-0559-9111-8	Sequence 8, Appl1
23	110.5	5.7 8257 4	US-09-054-9112-30	Sequence 30, Appl1
25	110	5.7 8257 4	US-08-242-932-8	Sequence 8, Appl1
25	110	5.7 1176 4	US-09-107-1320-659	Sequence 659, Appl1
28	109	5.6 1176 4	US-09-107-1320-659	Sequence 659, Appl1
28	109	5.6 5053 4	US-09-695-576-2277	Sequence 2277, Appl1
29	109	5.6 9391 4	US-08-932-993A-1	Sequence 2, Appl1
30	109	5.6 4200 4	US-09-248-796A-2277	Sequence 2277, Appl1
32	108.5	5.6 1536 4	US-08-946-028-58	Sequence 58, Appl1
32	108.5	5.6 1497 4	US-09-562-702A-1	Sequence 1, Appl1
34	108.5	5.6 2169 4	US-09-248-796A-6510	Sequence 6510, Appl1
35	108.5	5.6 9391 4	US-08-932-993A-869	Sequence 869, Appl1
35	108.5	5.6 3452 4	US-09-293-218	Sequence 293, Appl1
37	108.5	5.6 3607 4	US-09-710-379-4297	Sequence 4297, Appl1
38	108	5.6 3607 4	US-09-710-379-4297	Sequence 4297, Appl1
40	108	5.6 4283 4	US-09-293-5218	Sequence 5218, Appl1
41	108	5.5 1266 4	US-09-248-796A-4956	Sequence 4956, Appl1
42	107	5.5 789 4	US-09-248-796A-2	Sequence 2, Appl1
44	107	5.5 9626 4	US-10-667-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-916-421B-1/c
Patent application US/08916421B
Patent No. 6503729
Patent No. 6503729

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
FILE REFERENCE: P8275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE:
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
SOFTWARE: Patent In version 3.1
SEQ ID NO 1: 1664976
LENGTH: 1664976
TYPE: DNA
FEATURE:
ORGANISM: Methanococcus jannaschii
NAME/KEY: misc_feature
LOCATION: (28622)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 13:11:01.; search time 660 seconds
329,741 Million cell updates/sec

Title: US-08-876-132-3
Perfect score: 1947
Sequence: 1 NFIAMPKEIRGKKLMKKY......SQELLRKLSFVKFKRNCNC 371

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Xgapop	10.0	Xgapext	0.5
Fgapop	6.0	Fgapext	7.0
DelGP	6.0	Delext	7.0

Searched: 4390206 seqs, 2559870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 40%
Listing first 45 summaries

Database:
N_Geneseq_16Dec04:*
geneseq19908:*
geneseq19908:*
geneseq20006:*
geneseq20005a:*
geneseq20001bs:*
geneseq2002as:*
geneseq2003as:*
geneseq2003bs:*
geneseq2004as:*
geneseq2004bs:*

Command line parameters:
-MODEL=frame_p2n.model; -DEV=uln
-Q=/GP12 /USPTO #ppol_b/US088761132/runat_14022005_125513_22754/app_query.fasta_1.559
-DB=Geneseq; -DOPT=-fasta; -MATRIX=/home/users/...
-LIST=45; -LOCATION=200; -THR_SCORE=pct -THR_MIN=0 -ALIGN=45
-MODEL=LOCAL -OUTMED=20; -Next=HDR -BIGMATCH=200000000
-USER=US088761132 RCG ; -FILTER=HUMAN0.cdi
-NO_MMAP -LARGEQUERY ; -NEO_SCORES=0 -WAIT -DBBLOCK=100 -ICPU=3
-CPU=OUTPUT=12 -SORT=SCORES -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -CGAPOP=10 -DELGP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB	ID	Description
0	1943	99.8	1660	2	AAV74275	Pantoea c
1	146.5	7.6	35515	2	AAX20322	Borrelia
2	140.5	7.2	31..	2	AAX20052	Borrelia
3	139.5	7.1	3138	4	AAB76102	Full leng
4	139.5	7.2	3138	4	AAH46300	Pseudoalt
5	130.5	6.7	110000	2	AAV21209_11	Continuation

RESULT 1
AAV74275 standard; DNA; 1660 BP.

Key	Location/Qualifiers
CDS	480..1595

AAV74275;

16-JUN-1999 (first entry)
30-DEC-1998; 98MO-US012945.

Pantoea citrea cryptic plasmid ps.

Cryptic plasmid; ps; higher temperature; growth; elimination;
mobilization; ds.

Pantoea citrea

(GENV / GENEXOR INT INC.

ALIGNMENTS

			Score	ID	Description
0	6	129.5	26776	2	AAX20254	Borrelia
0	6	129	7078	2	AB103681	Drosophil
0	6	127.5	1198	2	AB103689	Drosophil
0	5	127.5	43576	2	AA68609	Bacteriop
0	10	110000	2	AAX20248_03	Continuation 14 of
0	7	127	4601	2	AB109845	Drosophil
0	8	122	2562	2	AB109845	Drosophil
0	9	122	1449	2	AAA70251	Plasmodiu
0	8	121.5	3449	2	AAY70251	Plasmodiu
0	6	121.5	2726	2	ACA33577	Prokaryot
0	6	121	1762	2	ACA33577	Prokaryot
0	6	119	2492	2	AAQ14755	PMS2 gene
0	4	118.5	3471	2	AD845300	Bacterial
0	7	118	2471	2	AAW55390	Oligonucl
0	6	118	5418	2	ACA33459	Prokaryot
0	6	117	6016	2	AAH57390	Human mRe
0	5	117	9841	2	AB129042_2	Primary
0	6	117	3669	2	AAY11951_2	Rice gene
0	6	116.5	1919	2	ABK63759	Rat seque
0	7	116.5	4301	2	ABS52308	Toxicity
0	7	116.5	4301	2	AD85245	Primary
0	8	116.5	4301	2	ABS52308	Plasmodiu
0	6	116.5	4301	2	AD85245	Rat astro
0	6	116	2479	2	ABL28642	Plasmodiu
0	6	116	3027	2	ACN47087	Human mRe
0	6	116	3033	2	AAn42087	Human sKe
0	6	116	3951	2	AB109844	Prokaryot
0	6	116	4702	2	ACN47087	Human mRe
0	6	116	89378	2	ADM46665_20	Continuation
0	6	116	89378	2	ADM46665_20	Continuation (21 o
0	6	110000	2	ADM46645_13	Continuation (20 o
0	6	110000	2	ADM46645_13	Continuation (20 o
0	6	110000	2	ADM7591	Thermococ
0	6	110000	2	ADN7209_19	Thermococ
0	6	110000	2	ADM46664_19	Continuation (20 o
0	6	110000	2	ADN7209_20	Thermococ
0	6	110000	2	ADM7960_20	Continuation

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: February 15, 2005, 13:21:06; Search time 5431 Seconds
(without alignments)
3308.832 Million cell updates/sec

ON protein - nucleic search, using frame plus p2n model

Title: 1 NFLHKVFKIRQKLNMKKY........SQELRLGSFVKRWICNMC 371
Perfect score: 1947
Sequence: US-08-876-132-3

Scoring table: BLOSUM62
   XgapOp 10.0   XgapExt   0.5
   NgapOp 6.0   NgapExt   0.5
   PgapOp 6.0   PgapExt   0.5
   DelOp   6.0   DelExt   7.0

Searched: 4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match Ot%
                 Maximum Match 100%
                 Listing first 45 summaries

Database:

Genbank:*
  1:   gb_htg:*
  2:   gb_htg:*
  3:   gb_in:*
  4:   gb_mat:*
  5:   gb_ov:*
  6:   gb_pat:*
  7:   gb_pln:*
  8:   gb_pln:*
  9:   gb_pri:*
 10:   gb_pri:*
 11:   gb_ro:*
 12:   gb_sts:*
 13:   gb_syn:*
 14:   gb_vi:*

Command line parameters:
  -MODEL=frame.p2n.model -DEV.xlh
  -Q=/CGI2_1/USPTO_Spool_3/US0887613Z/runat_14022005_125914_27172/app_Query.fasta_1-519
  -DB=Genbank -MAXSIZE=540 -MINSIZE=0 -ALIGN=50 -SLIST=45
  -UNITS=bits -START=1 -END=-1 -MATRIX=human1.col -LIST=45
  -DOCALIGN=20 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
  -OUTPT=pct -NOQuery -MAXNLIST=50 -MAXLINE=2000000000
  -USBR=USBR84613Z -DB=gen.1.1.8731.arunat_14022005_12914_27172_mCDfc6 -ICPU=3
  -NO_MAP -LARGEQUERY -NEG_SCORES=n -WAIT -DSPBLOCK=100 -LONGLOG
  -PGAFEXT=10 -PGAFOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

SUMMARIES
Result
No.   Score Match Length DB ID   Description
----
  1   170.5 9.8 11880   1   AF128899   Pantoea c
  2   170.5 9.8 11880   1   PFMA1P04   Pantoea c
  3   164.5 8.4 30299   1   AE001577   Borrelia
  4   155.5 8.0 3871   1   BBBGRCBCB   X87202 Borrelia bu

ALIGNMENTS

RESULT 1
LOCUS   AF128899   3661 bp   DNA   circular BCT 22-MAR-1999
DEFINITION Pantoea citrea strain ATCC31623 plasmid pPS0500, complete plasmid
            sequence
ACCESSION   AF128899
VERSION   AF128899.1 GI:4457229
KEYWORDS   .
SOURCE   Pantoea citrea
  ORGANISM Pantoea citrea
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1 (bases 1 to 3661)
  AUTHORS   Bilic-Nezic,M. and Delic,V.
  TITLE   Isolation and characterization of a cryptic plasmid from Erwinia
            citreus ATCC 3123
  JOURNAL   Plasmid. Microbiol. 83 (4), 485-492 (1997)
  MEDLINE   95012514
   PUBMED   9012014
REFERENCE   2 (bases 1 to 3661)
  AUTHORS   Bilic-Nezic,M. and Delic,V.
  TITLE   Sequence Analysis and Functional Characterization of Replication
            Region in plasmid pPS0500 from Pantoea citrea ATCC 3123
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 3661)
  AUTHORS   Bilic-Nezic,M. and Delic,V.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-FEB-1999) Biotechnology, PLIVA d.d., Research

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 15, 2005, 17:11:45 .. Search time 4043 Seconds
                                       3492.912 Million cell updates/sec

Title: us-08-876-132-3
Perfect_Score: 1947
Sequence: 1 NFLMWFEKIRKGHKLKMKKY......SQELLARGLSFVKMYKZKMC 371

Scoring table:
        BLOSUM62
        Xgapop 10.0 , Xgapax 0.5
        Ygapop 10.0 , Ygapax 0.5
        Fgapop 9.0 , Fgapax 0.5
        Delop   6.0 , Delex   7.0

Searched: 34239544 seqs, 1902134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
        -OMmodel frame_plus_p2n -DBtype ...

Database : EST...

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score from the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Q Match	Length	DB ID	Description
10	125.5	6.4	2740	3	AK029960 Mus muscu
11	124		571		CE766166 tigr-gss-
12	123.5	6.3	572		CE762696 tigr-gss-
13	123.5	6.3	672		CE600159 tigr-gss-
14	123.5	6.3	786		CE577943 tigr-gss-
15	123	6.3	786		CE577943 tigr-gss-
16	123	6.3	598		CE591938 tigr-gss-
17	123	6.3	621		CE591938 tigr-gss-
18	123	6.3	621		CB718400 tigr-gss-
19	123	6.3	709		CB718400 tigr-gss-
20	123	6.3	699		CE568932 tigr-gss-
21	123	6.3	699		CE509864 tigr-gss-
22	123	6.3	712		CE250864 tigr-gss-
23	123	6.3	755		CE176059 tigr-gss-
24	122.5	6.3	646		CB010297 tigr-gss-
25	122.5	6.3	742		CE038597 tigr-gss-
26	122.5	6.3	742		CE376783 tigr-gss-
27	122.5	6.3	669		CE376783 tigr-gss-
28	122.5	6.3	696		CE083270 tigr-gss-
29	123	6.3	673		CE147044 tigr-gss-
30	122.5	6.3	713		CE147044 tigr-gss-
31	122.5	6.3	720		CE196000 tigr-gss-
32	122.5	6.3	734		CE018890 tigr-gss-
33	122.5	6.3	734		CE759049 tigr-gss-
34	122.5	6.3	749		CE431832 tigr-gss-
35	122.5	6.3	749		CE808878 tigr-gss-
36	122.5	6.3	742		CE439065 tigr-gss-
37	122.5	6.3	696		CE083270 tigr-gss-
38	122	6.3	696		CE083270 tigr-gss-
39	122	6.3	713		CE596009 tigr-gss-
40	122	6.3	773		CE635080 tigr-gss-
41	122	6.3	818		CE103096 tigr-gss-
42	122	6.3	818		CE103096 tigr-gss-
43	122	6.2	688		CE778946 tigr-gss-
44	121	6.2	692		CB469175 tigr-gss-
45	121	6.2	692		CB469175 tigr-gss-

ALIGNMENTS

RESULT 1
CE673882
CE673882 -dog-170003294477931 Dog Library Canis familiaris genomic,

LOCUS   CE673882   711 bp   DNA   linear   GSS 29-SEP-2003
DEFINITION dog genomic survey sequence
ACCESSION   CE673882
VERSION   CE673882.1 GI:36992882
KEYWORDS   GSS.
SOURCE   Canis familiaris (dog)
  ORGANISM Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Caniformia; Canidae; Canis.
REFERENCE   1 (bases 1 to 711)
  AUTHORS   Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
  TITLE   The dog genome: survey sequencing and comparative analysis
  JOURNAL   Science 301 (5641), 1898-1903 (2003)
   PUBMED   14512627
COMMENT   Contact: Kirkness EF
            Genomic Research
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0208
            Fax: 301-838-0208
            Email: ekirkness@tigr.org
            Class: shotgun
FEATURES   Location/Qualifiers
     source   1..711
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"

```
                        GenCore version 5.1.6
           Copyright (c) 1993 - 2005 Compugen Ltd.

ON nucleic - nucleic search, using sw model

Run on:       February 15, 2005, 09:55:11
                        Search time 1230.12 seconds
                        without alignments
                        8446.087 Million cell updates/sec

Title:        US-08-876-132-2
Perfect score: 1847
Sequence:     1 AGATCTCAACCAGTTAAAA.........AACAACAAAACAGAATCT 1847

Scoring table: IDENTITY_NUC
               Gapop 10.0, Gapext 1.0

Searched:     5376673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Maximum DB seq length: 2000000000
Minimum DB seq length: 0

Post-processing: Minimum Match 0%
                 Listing first 45 summaries

Database :     Published Applications NA.*
               /cgm2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US10_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US10F_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US10_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US08_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US09_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US08C_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US09_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US81_PUBCOMB.seq.*
               /cgm2_6/prodata/1/pubpna/US81_NEW_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US81_PUB.seq.*
               /cgm2_6/prodata/1/pubpna/US6C_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score as good as or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.
```

```
c    0    12      96.2    5.1      6419    15   US-10-101-455-240    Sequence 240, App
c    0    15      94      5.1      8055    18   US-10-473-126-240    Sequence 306, App
c    0    15      91.4    5.0      6151    17   US-10-221-613-69     Sequence 69, App1
c    0    15      91.4    4.9      6292    17   US-10-221-714A-461   Sequence 461, App
c    0    15      90.2    4.9      8056    18   US-10-473-126-306    Sequence 306, App
c    0    15      89.6    4.9      6048    18   US-10-473-126-366    Sequence 366, App
c    0    15      89.6    4.9      6048    18   US-10-433-799-32     Sequence 32, Appl
c    0    15      89.6    4.9      6106    17   US-10-257-156-1445   Sequence 1445, Ap
c    0    15      89.6    4.9      6106    17   US-10-221-714A-151   Sequence 151, App
c    0    15      88.4    4.9      1300    18   US-10-473-126-741    Sequence 741, App1
c    0    15      88.4    4.9      1043    18   US-10-473-126-128    Sequence 128, App
c    0    15      88      4.8      1043    18   US-10-221-714A-96    Sequence 96, Appl
c    0    15      86.2    4.7      9760    17   US-10-221-613-1113   Sequence 1113, Ap
c    0    15      86.2    4.7      3683    18   US-10-473-126-339    Sequence 339, App
c    0    15      86      4.7      3693    18   US-10-473-126-1280   Sequence 1280, Ap
c    0    15      85.8    4.6      6056    18   US-10-311-455-1280   Sequence 1280, Ap
c    0    15      85.8    4.6      6048    18   US-10-473-126-240    Sequence 240, App
c    0    15      85.8    4.6      10048   17   US-10-221-613-211    Sequence 211, App
c    0    15      84.6    4.5      631     15   US-10-221-714A-339   Sequence 339, App
c    0    15      84.6    4.5      631     15   US-10-111-455-1460   Sequence 1460, Ap
c    0    15      82      4.5      3693    18   US-10-473-126-5746   Sequence 5746, Ap
c    0    15      82      4.5      5689    15   US-10-741-960-17905  Sequence 17905, A
c    0    37      81.6    4.4      49979   18   US-10-741-960-100    Sequence 100, App1
c    0    40      81.6    4.4      5689    15   US-10-221-613-100    Sequence 100, App1
c    0    40      81.6    4.4      11081   18   US-10-255-135-1675   Sequence 1675, Ap
c    0    43      81.4    4.4      8961    15   US-10-240-145-302    Sequence 302, App
c    0    44      81      4.4      8961    15   US-10-240-145-302    Sequence 302, App
c    0    45      81      4.4      9642    15   US-10-435-330     Sequence 330, App
```

ALIGNMENTS

```
RESULT 1
US-10-111-455-986/C
; Sequence 986, Application US/10311455
; Publication No. US20030104360A1
; GENERAL INFORMATION:
US-10-111-455-986
   APPLICANT: OLEK, Alexander
   APPLICANT: PIEPENBROCK, Christian
   APPLICANT: Rost
   TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detern
   TITLE OF INVENTION: cytosine methylation
   TITLE OF INVENTION: Prognosis of Diseases
   CURRENT FILING DATE:  2002-12-16
   CURRENT FILING DATE:  2003-07-02
   PRIOR FILING DATE:  2003-07-02 BP01/07537
   PRIOR APPLICATION NUMBER DE 10032539.7
   PRIOR APPLICATION NUMBER DE 100 32539.7
   PRIOR APPLICATION NUMBER:  DE 10043826.1
   PRIOR FILING DATE:  2000-09-01
   SEQ ID NOS: 2424
   SEQ ID NO: 986
   LENGTH: 7597
   TYPE: DNA
   ORGANISM: Artificial Sequence
   FEATURE:
   OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-111-455-986
   Query Match             7.0%;  Score 130;  DB 15;  Length 7597;
   Best Local Similarity   47.7%;
   Matches  447;  Conservative   0;  Mismatches  485;  Indels   6;  Gaps    2;

Qy  912  ATACCAGAGCCAAAACAAAACAAATAAACAACAAAATTTTCGACCGAAAAAATTTTT  971
```

```
                        GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.

ON nucleic    : nucleic search, using sw model

Run on:       February 15, 2005, 03:19:09 ; Search time 183.89 Seconds
                                              9628.26 Million cell updates/sec

Title:        US-08-876-132-2
Perfect Score: 1847
Sequence:     1 AGAYCTCAACCAGTTTAAA............AACCAACGAAAAGAGATCT 1847

Scoring table: IDENTITY NUC
              Gapop 10.0 , Gapext 1.0

Searched:     1202784 seqs; 818138159 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
                 Maximum Match 100%

Database :    Issued_Patents_NA:*
              /gns2_6/prodata/i/ina/5A_COMB.seq:*
              /gns2_6/prodata/i/ina/5b_COMB.seq:*
              /gns2_6/prodata/i/ina/6A_COMB.seq:*
              /gns2_6/prodata/i/ina/6B_COMB.seq:*
              /gns2_6/prodata/i/ina/PCTUS_COMB.seq:*
              /gns2_6/prodata/i/ina/backfiles1.seq:*

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

                              SUMMARIES

```
Result                Query
No.     Score   Match Length DB ID                    Description
 1      97.8    5.1   147382  US-09-949-016-14624      Sequence 14624, A
 2      88.6    4.8   187169  US-09-949-016-17067      Sequence 17067, A
 3      88.6    4.8   134987  US-09-949-016-13396      Sequence 13, Appl
 4      86.6    4.8   134987  US-09-949-016-11975      Sequence 11975, A
 5      86.6    4.7   134987  US-09-949-016-11971      Sequence 11971, A
 6      84.8    4.6   134987  US-09-949-016-13548      Sequence 13548, A
 7      84.8    4.6   134987  US-09-949-016-15349      Sequence 15349, A
 8      84.8    4.6   134987  US-09-949-016-15350      Sequence 15350, A
 9      84.8    4.6   134987  US-09-949-016-15350      Sequence 15350, A
10      84.8    4.6   134987  US-09-949-016-15508      Sequence 15508, A
11      82.6    4.5   134997  US-09-949-016-15509      Sequence 15509, A
12      82.6    4.5   194084  US-09-949-016-14231      Sequence 14231, A
13      81.8    4.4   63016   US-09-949-016-12423      Sequence 12423, A
14      80.6    4.4   18973   US-09-949-016-15852      Sequence 15852, A
15      80      4.4   205004  US-09-949-016-15851      Sequence 15851, A
16      80      4.4   205004  US-09-949-016-12387      Sequence 12387, A
17      80      4.2   223471  US-09-949-016-12724      Sequence 12724, A
18      80      4.2   223471  US-09-949-016-12724      Sequence 12724, A
19      77.6    4.2   18973   US-09-949-016-14164      Sequence 14164, A
20      77.6    4.2   612     US-09-949-016-14164      Sequence 14164, A
21      75.4    4.2   612     US-09-902-540-1357       Sequence 1357, Ap
22      75      4.1   612     US-09-949-016-01081      Sequence 1080, Ap
23      75      4.1   601     US-09-949-016-01080      Sequence 1080, Ap
24      75      4.1   601     US-09-949-016-37150      Sequence 37150, A
25      75      4.1   601     US-09-949-016-37151      Sequence 37151, A
26      75      4.1   601     US-09-949-016-15508      Sequence 15508,
27      75      4.1   601     US-09-949-016-14613G      Sequence 14613G,
```

                          ALIGNMENTS

```
RESULT 1.
US-09-949-016-14624/c
Sequence 14624 Patent Application US/0949016
Patent No. 681233
GENERAL INFORMATION:
  APPLICANT:... Craig et al
  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
  TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
  CURRENT APPLICATION DATA:
  CURRENT APPLICATION NUMBER: US/09/949,016
  CURRENT FILING DATE: 2000-10-14
  PRIOR APPLICATION NUMBER: 60/241,795
  PRIOR FILING DATE: 2000-10-20
  PRIOR APPLICATION NUMBER: 60/237,768
  PRIOR APPLICATION NUMBER: 60/231,498
  PRIOR FILING DATE: 2000-09-08
  OTHER INFORMATION: n = A,T,C or G
  SOFTWARE: PatentIn Version 4.0
SEQ ID NO 14624
LENGTH: 147382
TYPE: DNA
ORGANISM: Human
NAME/KEY: misc_feature
LOCATION: (1)..(147382)
SEQ ID NO: 14624

Query Match    Similarity    0.43'; Score 97.8; DB #;  Length 147382;
Best Local      Similarity    48.3%;
Matches  328'; Conservative   0; Mismatches  347; Indels  2; Gaps  2;
```

```
Qy    1005   TTTCCATTAATTAAAACCGAGATCTAAAACTAAAGATAYTAAGCAATCCGTTAGATTAT 1064
             |||||| | ||||||||| |||| ||||||||||||||||| ||||
Db   137331   TTCACGAAGTTTAAAACACAGTTAAACTTCTCTCTCTAAACTAAAGATAAGATTAT 137272

Qy    1065   AAAGTAGCAATAAAACAGTAAAATAACTATAATATTATTATYTTGATTATTATTTC 1124
             |||||| |||||||| ||| ||||||||||| |||||| ||||||
Db   137271   ATATAAATAATAATAAAAAATATAAATAATTATAATTAAATTAAAATATATTA 137212

Qy    1125   ACACTGATTATTTAAATATTAGAGATTATATTACTAAGTTAAACTAACTAAAAAAG 1184
             ||||| |||||||||| ||| |||| |||||||||| |||| ||||||||||
Db   137211   ATTATAAATATTAAAAAATTAAATATATTAATTATAAATAATAAATTAAAAG 137152

Qy    1185   AGGTAATAYTTTCAGTGTTATTTTAAAAACATTAGAAATAGAATGAACATTTATA 1244
             |||| ||| ||||||||| |||| || ||| ||| |||| |||| |||||
Db   137151   ATAATATATTATTAAACAGYATCATAAATAATAATAATAATATAATATAAAAGC 137092

Qy    1245   AAAGAGGGTATATCAATACTTTCTAACAATACTTACTTGATATAGATTAAAGC 1304
```

CN nucleic search, using sw model

GenCore version 5.1.6

Run on: February 15, 2005, 00:27:38 ; Search time 1003.82 seconds
                                      10892 200 Million cell updates/sec

Title: us-08-876-132-2
Perfect score: 1847
Sequence: 1 AGAYCCAACCAGTTAAA........AAGCAAGCAAAAGAGATCT 1847

Scoring table: IDENTIFY NUC
              Gapop 10.0 , Gapext 1.0

Searched: 4350206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
          1: geneseq01990s:**
          2: geneseq01990s:**
          3: geneseq02000s:*
          4: geneseq02010s:*
          5: geneseq02020as:*
          6: geneseq02030as:*
          7: geneseq02040b:*
          8: geneseq02050as:*
          9: geneseq02030bu:*
         10: geneseq02040as:*
         11: geneseq20030as:*
         12: geneseq20040as:*
         13: geneseq2004bs:*

SUMMARIES

Result No.	Score	Query Match Length	DB ID		Description
1	100.0	1847	1847	2	AAV74276 Pantoea c
2	7.0	2597	5		AB133013 Human imm
3	6.2	6109	6		AB032326 Human imm
4	6.2	6109	4		AB032326 Human gen
5	6.2	6109	5		AAS61077 Human gen
6	6.0	1423	8		AAS61077 Tumour su
7	6.0	1423	9		ABJ43158 Human su
8	6.0	1247	6		ACF62794 Colon can
9	5.3	8222	8		AAS65200 Colon can
10	5.3	8222	4		ACF62816 Colon can
11	5.3	11222	10		ADB54318 Pretreate
12	5.3	11222	10		ADB54318 Pretreate
13	5.3	20420	4		AAK63931 Human can
14	5.3	11331	4		AAS64621 Tumour su
15	5.3	19000	12		AB032659 Human can
16	5.2	15801	13		ADL17884 Human can
17	5.2	5413	4		AAS46694 Tumour su
18	5.2	5413	4		AAK69166 Human pho
19	5.2	19965	4		AAK69932 Human sec
20	5.2	6419	5		ABJ32267 Human imm

Pred. No is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

94.2	5.1	11521	8	ACA64845	Human HNR
94	5.1	8056	8	AB212246	Haematopo
93.4	4.9	46551	8	AB311201	Human opt
91.4	4.9	6154	8	ABX31201	Signal tr
91.4	4.9	6154	8	ABL70167	Chemicall
91.2	4.9	6393	4	Aas61112	Human gen
91.2	4.9	6393	4	Aas66116	Tumour su
90.2	4.9	7442	4	Aas46696	Human chr
90.1	4.9	5286	13	ADL10017-0	???
89.6	4.9	6648	6	AAS89279	Oligonucl
89.6	4.9	6648	6	ADG67002	Human apo
89.6	4.9	6106	13	ADG64429	Human ang
89.6	4.9	6106	13	ADG64429	Tumour su
88.8	4.8	5286	13	ABJ33472	Human inm
88.8	4.8	1544	8	ABJ34072	Human inm
88.8	4.8	1544	8	ABJ34155	Human inm
88.8	4.8	1000	12	ADG62833	Homopoly-
88.2	4.8	1000	12	ADG64711	Homopoly-
88.2	4.8	1300	12	AB091242	Synthetic
87.6	4.8	1300	12	Continuation (7 of
87.4	4.7	110000	12	ADB32968	6
87.4	4.7	110000	10	ABJ01242	Human inm
86.2	4.7	9766	4	AB091242	Signal pe
86.2	4.7	9766	4	ABL70197	Chemicall
45	4.7	9766	6	Aas61155	Human gen

RESULT 1
AAV74276
AAV74276 (standard; DNA; 1847 BP.

AAV74276;

16-JUN-1999 (first entry)

16-JUN-1999

Pantoea citrea cryptic plasmid pS.

Cryptic plasmid; pS; higher temperature; growth; elimination;
mobilization; de.

Pantoea citrea.

Pantoea citrea.

WO9859054-A1;

30-DEC-1998;

22-JUN-1998;

23-JUN-1997; 9708-00876132.

(GENW) GENENCOR INT INC.

Fowler TJ, Causey SC;

WPI; 1999-10562S/09.

Method for preparing an improved Enterobacteriaceae strain - useful for
improving bacterial fermentation strains of the family
Enterobacteriaceae.

Claim 19; Fig 1D-F; 30pp; English.

The sequence is that of one half of the cryptic plasmid pS from Pantoea
citrea. It can be used as part of a method of eliminating or reducing
plasmid from an Enterobacteriaceae strain within a method for reducing
mobilization of plasmid-encoded residing within the Enterobacteriaceae
strain. The elimination of the cryptic plasmid from the Enterobacteriaceae
permits growth of the organism at a higher temperature, which decreases
the production of the cryptic plasmid in the carbohydrate pathway. It
also has the commercial benefit of reducing both capital cost and

GenCore version 5.1.6

ON nucleic - nucleic search, using sw model

Run on: February 15, 2005, 00:33:13 /, Search time 7974.28 seconds
          11220.366 Million cell updates/sec

Title: US-08-876-132-2
Perfect_score: 1847
Sequence: 1 AGATCCAACAGTTAAAA.........AACAACGAAAACAGATCTT 1847

Scoring table: IDENTITY_NUC
                Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database : GenBank+

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length nb	ID	Description
1	1066.		3661	1	AF128989	Pantoea c
2	123.2		636	5	AX251345	Sequence
3	123.0		636	5	AX251345	Sequence
4	114.2		6109	6	AX251772	Sequence
5	114.2		6109	6	AX251772	Sequence
6	112.6		11861	9	AC069435	Sequence
7	112.6		9814	6	AX251284	Sequence
8	112.6		9814	6	AX251284	Sequence
9	112.6		127384	9	AC251284	Sequence
10	112.6		19981	1	AC009211	Drosophil
11	112.6		171105	3	AC007532	Drosophil
12	112.6		22298	1	AC003603	Drosophil
13	112.6		1237	1	AX347260	Sequence
14	112.6		170627	2	AC125567	Rattus no
15	112.4		250022	1	BX297091	Sequence
16	112.0		91709	1	AE014824	Homo sapi
17	110.		69970	3	FPMALLP3	Plasmodiu
18	109.		14867	1	AE001399	Homo sapi
19	108.8		90550	9	AL592166	Human DNA

ALIGNMENTS

RESUM.1
AF128989

LOCUS   AF128989   3661 bp   DNA   circular BCT 22-MAR-1999
DEFINITION Pantoea citrea strain ATCC1623 plasmid pPZG500, complete plasmid
            sequence.
ACCESSION   AF128989
VERSION   AF128989.1 GI:4457229
KEYWORDS   .
SOURCE   Pantoea citrea
  ORGANISM Pantoea citrea
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1 (bases 1 to 3661)
  AUTHORS   Bilic,M. and Delic,V.
  TITLE   Isolation and characterization of a cryptic plasmid from Erwinia
            citrea
  JOURNAL   J. Appl. Microbiol. 83 (4), 485-492 (1997)
   MEDLINE   98012514
REFERENCE   2 (bases 1 to 3661)
  AUTHORS   Bilic Nesic,M. and Delic,V.
  TITLE   Sequence analysis and functional characterization of replication
            region of plasmid pPZG500 from Pantoea citrea ATCC 3123
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 3661)
  AUTHORS   Bilic Nesic,M. and Delic,V.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-FEB-1999) Biotechnology, PLIVA d.d., Research
            Institute, Prilaz baruna Filipovica 25, Zagreb 10 000, Croatia
FEATURES   Location/Qualifiers
     source   1..3661
                     /organism="Pantoea citrea"
                     /mol_type="genomic DNA"
                     /strain="ATCC 3123"
                     /db_xref="ATCC:3123"
                     /db_xref="taxon:53336"
     misc_RNA   complement(331..3188..361,1..177))
                     /note="pPZG500"
                     /plasmid=pPZG500"
     CDS   99..317
                     /codon_start=1

20	108.8	5.9	210548	2	CR751608	Danio rer
21	108.7	5.9	128934	2	AC117629	Homo sapi
22		5.8		7	AC117742	Rattus no
23	105.2	5.8	121056	6	AC101990	Homo sapi
24	105.2	5.8		6	AC103590	Rattus no
25	105.2	5.7	145650	2	CR762412	Danio rer
26	105.2	5.7	141975	9	CR762412	Danio rer
27	105.2	5.7	161078	9	AP003500	Homo rer
28	104.4	5.7	161296	1	AC025120	Homo sapi
29	104.4	5.7	139010	2	AP003500	Homo rer
30	104.4	5.7	144772	1	CR382291	Danio rer
31	104.4	5.7	176700	7	CR382291	Danio rer
32	104.4	5.6	165955	2	CR792429	Homo rer
33	103.4	5.6	169555	9	CR792429	Danio rer
34	103.4	5.6	100925	3	CR782291	Homo rer
35	103.4	5.6	151628	7	AC012627	Homo sapi
36	103.4	5.6	135121	4	AC069925	Homo rer
37	103.2	5.6	160667	3	AC159025	Danio rer
38	103.2	5.6	142965	2	AC151623	Bos tauru
39	103.2	5.6	79038	9	AU627107	Homo sapi
40	103.	5.6	115357	2	AC762390	Danio rer
41	103.	5.6	150203	9	AC103590	Homo sapi
42	102.4	5.5	89807	2	AC105201	Homo sapi
43	102.4	5.5	89807	9	CR836602	Danio rer
44	101.8	5.5	131692	9	AC102292	Human DNA
45	101.8	5.5	131692	9	AL672277	Human DNA

ON nucleic - nucleic search, using sw method

Run on: February 15, 2005, 00:53:24 / Search time 6187.74 seconds
            11161.937 Million cell updates/sec

Title: US-08-876-132-2
Perfect score: 1847
Sequence: 1 AGATCTCAACGTTAAA.......AACCAGCAAAAACGATCT 1847

Scoring table: IDENTITY_NUC
            Gapop 10.0 , Gapext 1.0

Searched: 34235544 seqs; 19012134700 residues

Total number of hits satisfying chosen parameters: 68479088

Post-processing: 100%

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   ESTs $
             Minimum Match 0%
             Maximum Match 100%
             Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	121.6	1348	CO749499	CO749499	P043-4-AO
2	114.8	6.2	CO749728	CO749728	P043-1-A.CO
3	114	6.2	CG26147	CH216-159	
4	112.8	6.1	AG040947	AG040947	
5	112.8	6.1	CL119201	CL119201	SBSL-7621
6	112.6	6.1	CC759003	CC759003	P052-4-CO
7	112.4	6.1	CF753980	CF753980	P052-4-AO
8	112.2	6.0	AG386981	AG386981	Mus muscu
9	107.8	5.9	CL101704	CL101704	Mus muscu
10	107.8	5.9	AL065706	AL065706	Drosophil
11	107.8	5.8	CO83840	CO83840	Drosophil
12	107.8	5.8	ISB1-2814	ISB1-2814	
13	107.6	5.8	CL461494	CL461494	SAIL-1148
14	106.4	5.7	CL118721	CL118721	ISB1-7621
15	106.2	5.7	ISB1-7228	ISB1-7228	
16	106.2	5.7	CO82000	CO82000	CH216-165
17	105.6	5.7	CL107839	CL107839	CH216-151
18	105.4	5.6	AG346348	AG346348	CH216-49A
19	104.6	5.6	AD046318	AD046318	CH216-49A
20	104.2	5.6	CL400025	CL400025	CH216-167
21	104.2	5.6	CL007569	CL007569	CH216-167
22	104	5.6	CG732143	CG732143	111818090
23	104	5.6	CG732143	CG732143	11319460
24	103.8	5.6	CL016663	CL016663	CH216-1A2

25	103.6	5.6	CO749499	CO749499	P043-4-AO
26		5.6	AG386893	AG386893	Mus muscu
27		5.6	CL044158	CL044158	CH216-59P
28		5.6	CG744412	CG744412	P037-3-BO
29		5.6	CL081492	CL081492	
30		5.6	CO745119	CO745119	Mus muscu
31		5.6	CO745119	CO745119	P037-1-AO
32	102.4	5.6	CG40591	CG40591	11804530
33	102.2	5.6	CF753221	CF753221	P048-2-AO
34	102.2	5.5	CL081968	CL081968	CH216-165
35		5.5	CL081968	CL081968	CH216-165
36		5.5	CL509408	CL509408	SAIL_811
37		5.5	CO759757	CO759757	P037-1-PA
38	101.8	5.5	AG429511	AG429511	Mus muscu
39		5.5	CO759869	CO759869	P045-2-BI
40	101.6	5.5	AG429511	AG429511	Mus muscu
41		5.5	CO750869	CO750869	11804530
42		5.5	CL077121	CL077121	CH216-43
43		5.5	CL081212	CL081212	CH216-161
44		5.5	CO753083	CO753083	P048-1-CO
45		5.5	CL081995	CL081995	CH216-165
				CO753083	P048-1-CO

ALIGNMENTS

RESULT 1
CO749499/c
LOCUS   CO749499   1348 bp   DNA   linear GSS 24-OCT-2003
DEFINITION P043-4-AO6.za Ppa EcoRI BAC Library Pristionchus pacificus genomic.
ACCESSION   CO749499
VERSION   CO749499.1 GI:37970425
SSS
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1348)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            van de Meulen,H., and Sommer,R.
  TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (6), 715-722 (2003)
   PUBMED   12804007
   MEDLINE   22819551
COMMENT   Max-Planck Institute for Developmental Biology
            Evolutionary Biology
            Spemannstr. 37, 72076 Tuebingen D-72076, Germany
            Tel. 0049707161013;
            Fax. 0049707161013
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC_End
FEATURES   Location/Qualifiers
     source   1..1348
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /clone="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector"
ORIGIN

Qy   913 TACCGAGAACCAAACCAAAAAAAAATAAAAAATATAAAACAAAGAAAATTTTCCGCGAAAATATTTTC 972
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 TAATAAATAAAAATAAAAAAAAAAAAAAATTAAAAAATAATTAATTTTTAAAAATAAATAA 1018

Query Match   47.4%; Score 121.6; DB 9; Length 1348;
Best Local Similarity   47.4%;
Matches 427; Conservative 0; Mismatches 463; Indels 10; Gaps 2;

                 6.9/   Pred. No. 2.9e-13;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nuclide - nuclide search, using sw model

Run on: February 15, 2005, 09:55:11; Search time 1159.68 Seconds
8446.087 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAGGCGAATT........AAGGACTCGGAGCCAGTCT 1660

Scoring table: Gapop 10.0 , Gapext 1.0

Searched: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 5378673 seqs, 2951029984 residues

Minimum DB seq length: 1
Maximum DB seq length: 2000000000

Post-processing Minimum Match 0%
Maximum Match 90%
Listing first 45 summaries

Databases: Published Applications No.*
1: /cgm2.6/prodata/1/pubpna/US307_PUBCOMB.seq;*   10757146
2: /cgm2.6/prodata/1/pubpna/PCT_NEW_PUB.seq;*
3: /cgm2.6/prodata/1/pubpna/US806_NEW_PUB.seq;*
4: /cgm2.6/prodata/1/pubpna/US806_PUBCOMB.seq;*
5: /cgm2.6/prodata/1/pubpna/US807_PUBCOMB.seq;*
6: /cgm2.6/prodata/1/pubpna/PCT5_PUBCOMB.seq;*
7: /cgm2.6/prodata/1/pubpna/US5_NEW_PUB.seq;*
8: /cgm2.6/prodata/1/pubpna/US5_PUBCOMB.seq;*
9: /cgm2.6/prodata/1/pubpna/US05_NEW_PUB.seq;*
10: /cgm2.6/prodata/1/pubpna/US905_PUBCOMB.seq;*
11: /cgm2.6/prodata/1/pubpna/PCT_NEW_PUB.seq;*
12: /cgm2.6/prodata/1/pubpna/US90C_PUBCOMB.seq;*
13: /cgm2.6/prodata/1/pubpna/US10_PUBCOMB.seq;*
14: /cgm2.6/prodata/1/pubpna/US10C_PUBCOMB.seq;*
15: /cgm2.6/prodata/1/pubpna/US10_NEW_PUB.seq;*
16: /cgm2.6/prodata/1/pubpna/US10C_PUBCOMB.seq;*
17: /cgm2.6/prodata/1/pubpna/PCT5_PUBCOMB.seq;*
18: /cgm2.6/prodata/1/pubpna/US0_PUB_COMB.seq;*
19: /cgm2.6/prodata/1/pubpna/US1_PUB_COMB.seq;*
20: /cgm2.6/prodata/1/pubpna/US0_NEW_PUB.seq;*
21: /cgm2.6/prodata/1/pubpna/US0_PUBCOMB.seq;*
22: /cgm2.6/prodata/1/pubpna/US60_PUBCOMB.seq;*

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	121.8	7.3	1300	18	US-10-668-749A-1	Sequence 1, Appli
2	104.2	5.8	367	16	US-10-211-714A-461	Sequence 461, App
3	101.4	5.7	1243	18	US-10-425-115-112717	Sequence 112717
4	96.5	5.8	1803	18	US-10-471-126-139	Sequence 139, App
5	94.8	5.7	778	18	US-10-363-345A-2179	Sequence 2179, Ap
6	94.8	5.7	778	18	US-10-363-345A-2180	Sequence 2180, Ap
7	92.5	5.4	1297	18	US-10-425-115-67687	Sequence 67687, A
8	90.2	5.4	14006	15	US-10-311-455-1931	Sequence 1931, Ap

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.

RESULT 1
US-10-668-749A-1
; Sequence 1, Application US/10668749A
; Publication No. US20040101205A1
; GENERAL INFORMATION:
; APPLICANT: Agilent Technologies
; TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis
; FILE REFERENCE: 10012-1580
; CURRENT APPLICATION NUMBER: US/10/668,749A
; CURRENT FILING DATE: 2003-09-23
; SEQ ID NO 1
; SOFBM OF SEQ ID NOS: 8
; SOFBM ID: Patentln version 3.2
; TYPE: DNA
; LENGTH: 1300
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: synthetic construct
; FEATURE:
US-10-668-749A-1

Query Match 98.3%; Score 121.8; DB 18; Length 1300;
Best Local Similarity 44.3%; Pred. No. 2.8e-12;
Matches 495; Conservative 0; Mismatches 622; Indels 0; Gaps 0;

ALIGNMENTS

Query	7.3	Score 121.8;	DB 18;	Length 1300;	
Best Local Similarity 44.3%		Pred. No. 2.8e-12;			
Matches 495;	Conservative 0;	Mismatches 622;	Indels 0;	Gaps 0;	

CanCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

CM nuclid - nuclid search, using sw model

Run on: February 15, 2005, 03:19:00 / Search time 282.11 seconds
(without updates)
9628.226 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCAAATY ..........AAGACTCCAACCAGTACACCT 1660

Scoring table: IDENTIFY_NUC
Gapop 10.0 / Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Minimum Match length: 100%
Maximum DB seq length: listing first 45 summaries

Post-processing:

Databases: Issued Patent_NA: *
1: /cgm2_6/procdata/l/ina/na_COMB.seq;*
2: /cgm2_6/procdata/l/ina/in/na_COMB.seq;*
3: /cgm2_6/procdata/l/ina/na_COMB.seq;*
4: /cgm2_6/procdata/l/ina/na_COMB.seq;*
5: /cgm2_6/procdata/l/ina/na_COMB.seq;*
6: /cgm2_6/procdata/l/ina/PCTUS_COMB.seq;*
7: /cgm2_6/procdata/l/ina/hacfiles1.seq;*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Query Match Length	Score	DB	Pred. No.	DB ID	Description
1	612	93.6	5.7		US-09-902-540-1357	Sequence 1357, Ap
2	1039	82	4.9		US-09-902-540-1280	Sequence 1280, Ap
3	1039	72.3	4.8		US-09-806-708B-12	Sequence 14, Appl
4	1141	70.2	4.3		US-09-806-708B-13	Sequence 13, Appl
5	19124	70.2	4.1		US-08-047-8289-13	Sequence 13, Appl
6	5852	68.9	4.0		US-07-867-1082	Sequence 1318, Appl
7	16649	67.8	4.0		US-09-540-1318	Sequence 1318, Appl
8	1664976	64.4	3.9		US-08-099-552-510-1	Sequence 1, Appl
9	1956	64.4	3.9		US-09-092-552-510-1	Sequence 1, Appl
10	615	64.4	3.9		US-09-098-416-1386	Sequence 186, App
11	107976	62.9	3.7		US-09-949-016-13540	Sequence 12776, A
12	191567	60.6	3.7		US-09-949-016-15129	Sequence 15129, A
13	55986	60.5	3.6		US-09-949-016-15129	Sequence 15129, A
14	55986	60.4	3.6		US-09-949-016-998-1	Sequence 13, Appl
15	1956	60.2	3.6		US-09-551-998-1	Sequence 13, Appl
16	1956	60.2	3.6		US-09-551-998-1	Sequence 13, Appl
17	640681	60.2	3.6		US-09-790-088-1	Sequence 1, Appl
18	205044	59.6	3.4		US-09-949-016-15881	Sequence 15881, A
19	205044	59.4	3.4		US-09-949-016-15881	Sequence 15881, A
20	205044	59.4	3.4		US-09-949-016-15883	Sequence 15883, A
21	223471	59.4	3.4		US-09-949-016-12397	Sequence 12397, A
22	223471	59.4	3.4		US-09-949-016-12394	Sequence 12394, A
23	223471	59.4	3.4		US-09-949-016-12725	Sequence 12725, A
24	10643	58.9	3.5		US-09-447-495-1	Sequence 1, Appl
25	16649	58.8	3.5		US-09-640-173-53	Sequence 53, Appl
26	396	58.6	3.5		US-09-713-550-53	Sequence 53, Appl
27	396	58.6	3.5			Sequence 53, Appl

28	396	58.6	3.5		US-09-825-294-53	Sequence 53, Appl
29	259396	58.6	3.5		US-09-970-666-53	Sequence 53, Appl
30	2872	58.6	3.5		US-09-949-016-208	Sequence 208, App
32	837	58.1	3.5		US-09-949-016-10531	Sequence 10531, A
34	240	57.2	3.4		US-08-628-17-6	Sequence 6, Appl
35	74059	57.2	3.4		US-09-949-416-1785	Sequence 11785, A
36	2570	57.2	3.4		US-09-949-016-11137	Sequence 1137, Ap
37	601	57.2	3.4		US-09-949-016-12239	Sequence 2, Appl
38	601	57	3.3		US-09-949-016-15868	Sequence 6, Appl
39	601	57	3.3		US-09-949-016-146136	Sequence 146136, A
40	601	56.6	3.3		US-08-897-126-22	Sequence 22, Appl
42	601	56.6	3.3		US-09-800-729-33	Sequence 33, Appl
44	396	55.4	3.3			Sequence 33, Appl
45	2394	55.4	3.3			

ALIGNMENTS

RESULT 1
US-09-902-540-1357
Sequence 1357, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
  APPLICANT: Goldman, Barry S.
  APPLICANT: Hinkle, Gregory C.
  APPLICANT: Slater, Steven C.
  APPLICANT: Wasgand, Roger C.
  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
  NUMBER OF SEQ ID NOS: 16825
  CURRENT APPLICATION NUMBER: US/09/902,540
  CURRENT FILING DATE: 2001-07-10
  PRIOR APPLICATION NUMBER: 60/217,883
  PRIOR APPLICATION FILING DATE: 2000-07-10
  SEQ ID NO: 1357
  LENGTH: 612
  TYPE: DNA
  ORGANISM: Myxococcus xanthus
  FEATURE:
  NAME/KEY: unsure
  LOCATION: (1)..(612)
  OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match   5.7%; Score 93.8; DB 4; Length 612;
Best Local Similarity   47.1%;
Matches 284; Conservative 0; Mismatches 319; Indels 0; Gaps 0;

16 AANTCGAAAAAATAGATAAAATTCCCGCAGTTAAAACCGACTCTAAACAAATAGTCG 75
   ||||| ||||||||| || |||| || |||||||||||||| || |||||||||
76 AAGAGCAAAACCAAACAAAATAAATACATATTTCAGTAGTAGGAGAAGAAAAATA 135

76 AACAAGACTCGGCTTTACACTCCCAAACCAACAAATAAATAGAAAGAGACTGTGAT 135
   ||| ||||||| ||||||| |||||| ||||||||||||||||| |||||||
136 AAGAAGCTGGGCTCTCACTCCCAAAGGAAGAAATAAATAAAAGAAGACTCTAT 195

116 AAGAAGACTCGGCTTTACACTCCCAAACCAACAAATAAATAGAAAGAGACTGTGAT 195
   ||| ||||||| ||||||| |||||| ||||||||||||||||| |||||||
180 AAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAT 240

196 TTTTAATCGGAATCGTGCGAGAAAGAAAATTTATTTCATTTCGAGGATATATTC 255
   |||||||| ||| ||| | ||||||||| ||| ||| |||||| |||||||
240 ATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAATA 180

181 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAATTA 135

256 TGTAAGTTCATAAAAACTTGATAAAAAAAAATCACATCAAAAAATGTGTTATTTGACA 315
   | || || ||||| ||| ||||||| |||||||| |||||||||||||
300 AAAAATAAAAAAATAAAAAAAAATAAAAATTAAAAAAAAATTAAATAAAT 300

241 AAAAATAAAAAAAAAAAATAAATTAAAAAAAATAATT 300

GenCore version 5.1.6

OM nuclac - nuclaic search, using sw model

Run on: February 15, 2005, 00:27:38; Search time 902.184 Seconds
(without ...) 10892.200 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCAAATT.........AAGGATCCGACCCAGATCT 1660

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Post-processing: Minimum Match 018
listing first: 45 summaries

Minimum DB seq length: 10
Maximum DB seq length: 2000000000

Searched: 439206 seqs, 295987067 residues

Total number of hits satisfying chosen parameters: 8780412

Database: N_Geneseq_14Spe14.*
1: geneseq5q15904.*
2: geneseq119906:*
3: geneseqn20a00a0.*
4: geneseqn20a05a.*
5: geneseqn20l0bs.*
6: geneseqn20a02as.*
7: geneseqn20a03as.*
8: geneseqn20a00as.*
9: geneseqn20a03a.*
10: geneseqn20o03b.*
11: geneseqn20a04as.*
12: geneseqn20a04bs.*
13: gnenseqn20a0bs.*

Pred. No. is the number of results predicted by chance
to have a score greater than or equal to the score being printed,
and is derived by analysis of the local score distribution.

SUMMARIES

Result No.	Score	Match %	Query Match	Length	DB	ID	Description
1	1657	99.8	1660		12	AAV74275	Pantoea c
2	121	9.3	1291		8	AAB05917	Synthetic
3	109	6.5	6292		4	Adp65911	Synthetic
4	107.6	6.5	3681		10	AAd92827	Homopoly-
5	107.6	6.5	3681		12	AAd92833	Homopoly-
6	96.6	5.7	3683		8	AAd92832	Homopoly-
7	95	5.7	3683		8	AAd92832	Homopoly-
8	94.8	5.7	778		6	AAd10053	Haemactopo
9	94.8	5.7	778		6	ABZ10053	Haemacapo
10	93.6	5.6	6644		12	AAx33181	Base sequ
11	93.6	5.6	7372		2	AAx33182	Base sequ
12	93.6	5.6	7996		2	AAx33184	Base sequ
13	90.4	5.4	8222		8	ACf62816	Colon can
14	90.2	5.4	11329		6	ABd53958	Prostate
15	90.2	5.4	6666		8	ABj33697	Human imm
16	88.8	5.3	6666		8	Acf63797	Colon can
17	88.8	5.3	11222		10	AAb54130	Prostate
20	88.2	5.3	975		6	Abq29508	Oligonucl

21	88.2	5.3	975		6	Abq29509	Oligonucl
22	87.6	4.9	7442		8	AAd46686	Tumour su
23	83.8	5.0	9539		4	AAb45415	Human imm
24	83.8	5.0	9539		4	AAb45417	Human imm
25	83.8	5.0	9539		4	AAb45419	Human imm
26	83.8	4.0	7821		8	AAk28180	DNA trans
27	83.6	4.6	7823		8	ABj12527	Human imm
28	83.6	4.6	7823		8	AAk25127	Human imm
29	83.6	4.6	7823		8	ABj13061	Human imm
30	83.6	4.6	7823		8	ABj13061	Human imm
31	81.6	4.9	11745		8	AAk21493	Signal tr
32	80	4.8	7814		6	AAk46530	Tumour su
33	79.8	4.7	12237		4	AAk28832	DNA trans
34	78.4	4.7	15674		8	ABj34458	Human imm
35	78.4	4.7	15674		8	ABj23263	Human imm
36	78.4	4.7	2943		12	AAa70229	Chemical
37	78.4	4.7	5276		6	ABj23150	Human imm
38	78.4	4.7	15674		8	ABj70514	Human imm
40	78.4	4.7	15674		8	ABj70511	Human imm
42	77.4	4.7	15674		8	AAd99738	Complemen
43	77.4	4.7	5493		6	AAd93014	Human imm
44	77.4	4.7	8056		8	AB210246	Oligonucl
45	76.8	4.7	6314		6	AB154314	Chemical

ALIGNMENTS

RESULT 1
AAV74275

AAV74275 standard; DNA; 1660 BP.

XX
AC AAV74275;
XX
DT 16-JUN-1999 (first entry)
XX
DE Pantoea citrea cryptic plasmid pS.
XX
KW Cryptic plasmid; pS; higher temperature; growth; elimination;
KW mobilization of...
XX
OS Pantoea citrea.
XX
FH Key   Location/Qualifiers
FH
FT CDS   488..1595
FT   /note: "largest open reading frame"
XX
PN WO9859054-A1.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98NO-US012945.
XX
PR 23-JUN-1997; 97US-00878132.
XX
PA (GENV) GENENCOR INT INC.
XX
PI Fowler T, Causey SC;
XX
PT Method for preparing an improved Enterobacteriaceae strain - useful for
PT improving bacterial fermentation strains of the family
PT Enterobacteriaceae.
XX
PS Claim 18; Fig 1A-C; 30pp; English.
XX
CC The sequence is that of one half of the cryptic plasmid pS from Pantoea
CC citrea. It can be used as part of a method of eliminating the cryptic

GenCore version 5.1.6
Copyright (c) 1993 -2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2005, 00:33:38 / Search time 7168.72 Seconds
                 (without nucleic)
                 sequence/sec
                 1120.366 Million Cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCGAATT .... AAGGACTCCGAGCCAGATCT 1660

Scoring table: IDENTIFY_NUC
                 Gapop 10.0, Gapex 1.0

Searched: 4708233 segs, 2422760790 residues

Total number of hits satisfying chosen parameters: 941666

Minimum DB seq length: 1
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
                 Maximum Match 100%
                 Lieing first 45 summaries

Database :   Genbnbl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
1	1308.6	78.8	3661	6	AF128889	AF128889 Pantoea c
2	120.4	7.3	170627	6	AC125567	AC125567 Rattus no
3	120.4	7.3	170627	6	AC125567	AC125567 Rattus no
4	117.6	7.1	66993	6	AC138074	AC138074 Homo sapi
5	112.4	6.8	93791	6	AC118073	AC118073 Rattus no
6	109	6.6	6292	6	AX251493	AX251493 Sequence
7	104.2	6.3	349960	6	AC114558	AC114558 Plasmodiu
8	104	6.3	266208	6	AC117140	AC117140 Rattus no
9	104	6.3	266208	6	AC117140	AC117140 Rattus no
10	101.6	6.1	118882	6	ALF27227	Human DNA
11	101.6	6.1	231912	6	AC087566	AC087566 Homo sapi
12	99.6	6.0	174384	6	AC009524	AC009524 Homo sapi
13	98.8	5.9	130540	6	AX843153	Continuation (4 of
14	98.8	5.9	11000	6	AX679410_3	Continuation (4 of
15	96.6	5.8	3683	6	AX598999	AX598999 Sequence
16	96.6	5.8	3683	6	AX598999	AX598999 Sequence
17	96.6	5.8	84667	6	AE001398	AE001398 Danio rer
18	96.6	5.8	84667	6	AE001398	AE001398 Danio rer
19	96	5.8	3683	6	AX598053	AX598053 Sequence

20	95.4	5.7	83435	5	BX640669	BX640669 Zebrafish
21	93.8	5.7	175544	5	ACL17342	ACL17342 Rattus no
22	93.8	5.7	289993	5	ACL35678	ACL35678 Rattus no
23	93.6	5.6	349969	5	AC009524	AC009524 Mus muscu
24	93.6	5.6	6644	5	E23356	E23356 Virus vecto
25	93.6	5.6	7972	5	E23357	E23357 Virus vecto
26	93.6	5.6	7372	5	E23357	E23357 Virus vecto
27	93.6	5.6	7996	5	E23355	E23355 Virus vecto
28	93.2	5.6	105682	5	ACL16957_3	Continuation (4 of
29	91.8	5.5	67970	5	AC011746	AC011746 Plasmodiu
30	91.8	5.5	67970	5	AC011746	AC011746 Plasmodiu
31	91.2	5.5	104767	5	ACS11623	ACS11623 Sequence
32	91.2	5.5	25079	5	CRY11282A	CRY11282A Caenorha
33	91.0	5.5	131026	5	BX004843	BX004843 Zebrafish
34	90.8	5.4	11222	5	BX001361	BX001361 Danio rer
35	90.4	5.4	12222	5	AX822488	AX822488 Sequence
36	90.4	5.4	11222	5	AX826122	AX826122 Sequence
37	90.4	5.4	137951	5	AC346860	AC346860 Sequence
38	90.2	5.4	137951	5	CR936840	CR936840 Danio rer
39	90.2	5.4	257703	5	CR847802	CR847802 Danio rer
40	90.2	5.4	11222	5	BX005871	BX005871 Danio rer
41	90.2	5.4	11222	5	BX548072	BX548072 Danio rer
42	90.2	5.4	110030	5	AC094431	AC094431 Mus muscu
43	90	5.4	300695	5	ZA80853	ZA80853 Caenorhabd
44	89.8	5.4	300695	5	AC094431	AC094431 Mus muscu
45	89.8	5.4	6666	5	AX346599	AX346599 Sequence
46	87.8	5.4	110743	5	BX649189	BX649189 Zebrafish

ALIGNMENTS

RESULT 1
AF128889
LOCUS   AF128889   3661 bp   DNA   circular BCT 22-MAR-1999
DEFINITION Pantoea citrea strain ATCC1623 plasmid pP20500, complete plasmid
            sequence.
ACCESSION   AF128889
VERSION   AF128889.1 GI:4457229
KEYWORDS   .
SOURCE   Pantoea citrea
  ORGANISM Pantoea citrea
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1 (bases 1 to 3661)
  AUTHORS   Bilic,M. and Dело,V.
  TITLE   Isolation and Characterization of a cryptic plasmid from Erwinia
            (citrea ATCC 1623)
  JOURNAL   Plasmid 38 (4), 485-492 (1997)
  PUBMED   9801235
REFERENCE   2 (bases 1 to 3661)
  AUTHORS   Bilic-Nesic,M. and Delic,V.
  TITLE   Sequence Analysis and Functional Characterization of Replication
            Unpublished   plasmid pP20500 from Pantoea citrea
REFERENCE   3 (bases 1 to 3661)
  AUTHORS   Bilic-Nesic,M. and Delic,V.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-FEB-1999) Biotechnology, PLIVA d.d., Research
            Institute, Location/Qualifiers
FEATURES   Location/Qualifiers
     source   1..3661
                     /organism="Pantoea citrea"
                     /mol_type="genomic DNA"
                     /strain="ATCC1623"
                     /db_xref="taxon:3533c"
     misc_RNA   1..3661
                     /plasmid="pP20500"
                     complement(join(3188..3661.1..177))
     CDS   /note="ORT1"
                     99..317
                     /note="Ofi"
                     /codon_start1

OM nucleic - nucleic search, using sw model.

GenCore version 5.1.6
Copyright (c) 1993 -2005 Compugen Ltd.

Run on :   February 15, 2005, 00:53:24 ; Search time 556l.26 Seconds
                                                 1136l.937 Million cell updates/sec

Title: US-08-876-132-1
Perfect score: 1660
Sequence: 1 AGATCTACACAAGGCGAATT.........AAGGACCTCCGAGCCAGATCT 1660

Scoring table: Gapop 10.0 , Gapext 1.0
                IDENTITY_NUC

Searched: 3423554 seqs, 1903213124700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
                  Maximum Match 100%
                  Listing First 45 Summaries

Database :   EST_si (est.i:*
                1: gb_set.i:*
                2: gb_bct:*
                3: gb_est4:*
                4: gb_est3:*
                5: gb_est1:*
                6: gb_est2:*
                7: gb_est:*
                8: gb_gss.i:*
                9: gb_gss1:*
                9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147.2	1	1594	CL078613	CL078613 CH216-151
2	143	9	1407	AJ592026	AJ592026 Arabidops
3	141.4	9	1599	CL077957	CL077957 CH216-165
4	140.8	8	1599	CL083940	CL083940 Arabidops
5	140	8	1842	CL084158	CL084158 CH216-167
6	139.8	4	1491	CL041415	CL041415 CH216-59P
7	139.4	9	1491	CL075321	CL075321 P048-J-AO
8	137.9	8	1493	CL075322	CL075322 CH216-151
9	135.9	8	1693	CL078589	CL078589 CH216-151
10	135.5	8	1503	CL075144	CL075144 CH216-151
11	135.4	8.1	1536	CL079838	CL079838 P045-1-BO
12	134.9	8.1	1434	AJ592058	AJ592058 Arabidops
13	134.8	8.1	1434	AJ592058	AJ592058 Arabidops
14	134.5	9	1451	CL041325	CL041325 Mus muscu
15	134.4	8.0	1451	AJ591978	AJ591978 Arabidops
16	133.7	8.0	1334	AG341525	AG341525 Mus muscu
17	133.4	8.0	1334	AJ592058	AJ592058 Arabidops
18	133.4	8.1	1594	AG343229	AG343229 Mus muscu
19	133.2	9	1594	CL038406	CL038406 Mus muscu
20	131.8	9	1784	CL078634	CL078634 CH216-46A
21	131.7	7.9	1781	CS757757	CS757757 CH216-165
22	131.6	7.9	1784	CL081992	CL081992 CH216-165
23	131.4	7.9	1594	CL082020	CL082020 Forward s
24	130.9	7.9	1259	CR004722	CR004722 Forward s

(additional entries continue)

RESULT 1
CL078613
LOCUS   CL078613   1594 bp   DNA   linear
DEFINITION CH216-151MN17_Sg5.1_CH216 Xenopus tropicalis genomic clone.
ACCESSION CH216-151M17; genomic survey sequence.
VERSION   CL078613.1 GI:40534526
KEYWORDS   GSS.
SOURCE   Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
           Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1594)
AUTHORS   Mardis,E.R. and Wilson,R., McPherson,J., Warren,W., Graves,T.,
JOURNAL   Unpublished (2005)
TITLE   A physical map of the Xenopus tropicalis genome
COMMENT   Contact: Richard K Wilson
           Genome Sequencing Center
           Washington University School of Medicine
           Seq primer: SP6 accgcgcgttcgacct
           Insert length: 175000 Std Error: 0.00
           High quality sequence start: 557
           High quality sequence stop: 634.
           Local/Qualifiers
           /mol_type="genomic DNA"
           /organism="Xenopus tropicalis"
           /strain="Nigerian"
           /db_xref="taxon:8364"
           /sex="male"
FEATURES   source
           1594
           /cell_line="Stock 248 F7A2; inbred N7"
           /clone="CH216"
           /clone_lib="CH216-151M17"
           /note="Vector: pTARBAC2.1 CHORI-216 Xenopus tropicalis
           BAC library"
ORIGIN
           1 AGATCTACACAAGGCGAATT...

ALIGNMENTS

QY
Query Match   8.9%;   Score 147.2; DB bp   Length 1594;
Best Local Similarity 64.5%;   Pred. No. 5.2e-17;
Matches 500; Conservative 0; Mismatches 623; Indels 0; Gaps 0;